

Allen

16C!
8/18

1645ALLEN

#11/72

09/18/99

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/056,019ADATE: 07/07/1999
TIME: 14:28:17

Input Set: I056019A.RAW

This Raw Listing contains the General Information
Section and up to first 5 pages.

ENTERED

1 <110> APPLICANT: Tuomanen, Elaine I
2 Wizemann, Theresa
3 Masure, H. R.
4 Johnson, Leslie S.
5 Koenig, Scott
6 <120> TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
7 CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
8 THEREFROM AND USES THEREOF
9 <130> FILE REFERENCE: 1340-1-017
10 <140> CURRENT APPLICATION NUMBER: US/09/056,019A
11 <141> CURRENT FILING DATE: 1998-04-07
12 <160> NUMBER OF SEQ ID NOS: 39
13 <170> SOFTWARE: PatentIn Ver. 2.0
14 <210> SEQ ID NO 1
15 <211> LENGTH: 406
16 <212> TYPE: PRT
17 <213> ORGANISM: Streptococcus pneumoniae
18 <400> SEQUENCE: 1
19 Glu Asn Glu Gly Ala Thr Gln Val Pro Thr Ser Ser Asn Arg Ala Asn
20 1 5 10 15
21 Glu Ser Gln Ala Glu Gln Gly Glu Gln Pro Lys Lys Leu Asp Ser Glu
22 20 25 30
23 Arg Asp Lys Ala Arg Lys Glu Val Glu Glu Tyr Val Lys Lys Ile Val
24 35 40 45
25 Gly Glu Ser Tyr Ala Lys Ser Thr Lys Lys Arg His Thr Ile Thr Val
26 50 55 60
27 Ala Leu Val Asn Glu Leu Asn Asn Ile Lys Asn Glu Tyr Leu Asn Lys
28 65 70 75 80
29 Ile Val Glu Ser Thr Ser Glu Ser Gln Leu Gln Ile Leu Met Met Glu
30 85 90 95
31 Ser Arg Ser Lys Val Asp Glu Ala Val Ser Lys Phe Glu Lys Asp Ser
32 100 105 110
33 Ser Ser Ser Ser Ser Ser Asp Ser Ser Thr Lys Pro Glu Ala Ser Asp
34 115 120 125
35 Thr Ala Lys Pro Asn Lys Pro Thr Glu Pro Gly Glu Lys Val Ala Glu
36 130 135 140
37 Ala Lys Lys Lys Val Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys
38 145 150 155 160
39 Glu Glu Asp Arg Arg Asn Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Glu
40 165 170 175
41 Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala Glu Leu Glu
42 180 185 190
43 Leu Val Lys Val Lys Ala Asn Glu Pro Arg Asp Glu Gln Lys Ile Lys
44 195 200 205

PAGE: 2

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/056,019A

DATE: 07/07/1999
TIME: 14:28:17

Input Set: I056019A.RAW

```

45      Gln Ala Glu Ala Glu Val Glu Ser Lys Gln Ala Glu Ala Thr Arg Leu
46          210                215                220
47      Lys Lys Ile Lys Thr Asp Arg Glu Glu Ala Glu Glu Glu Ala Lys Arg
48          225                230                235                240
49      Arg Ala Asp Ala Lys Glu Gln Gly Lys Pro Lys Gly Arg Ala Lys Arg
50          245                250                255
51      Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala
52          260                265                270
53      Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Pro Ser Pro Ser
54          275                280                285
55      Leu Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu
56          290                295                300
57      Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr
58          305                310                315                320
59      Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp
60          325                330                335
61      Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys
62          340                345                350
63      Glu Pro Arg Asn Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu
64          355                360                365
65      Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg
66          370                375                380
67      Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu Asp Lys
68          385                390                395                400
69      Val Lys Glu Lys Pro Ala
70          405
71      <210> SEQ ID NO 2
72      <211> LENGTH: 655
73      <212> TYPE: PRT
74      <213> ORGANISM: Streptococcus pneumoniae
75      <400> SEQUENCE: 2
76      Glu Asn Glu Gly Ala Thr Gln Val Pro Thr Ser Ser Asn Arg Ala Asn
77          1                5                10                15
78      Glu Ser Gln Ala Glu Gln Gly Glu Gln Pro Lys Lys Leu Asp Ser Glu
79          20                25                30
80      Arg Asp Lys Ala Arg Lys Glu Val Glu Glu Tyr Val Lys Lys Ile Val
81          35                40                45
82      Gly Glu Ser Tyr Ala Lys Ser Thr Lys Lys Arg His Thr Ile Thr Val
83          50                55                60
84      Ala Leu Val Asn Glu Leu Asn Asn Ile Lys Asn Glu Tyr Leu Asn Lys
85          65                70                75                80
86      Ile Val Glu Ser Thr Ser Glu Ser Gln Leu Gln Ile Leu Met Met Glu
87          85                90                95
88      Ser Arg Ser Lys Val Asp Glu Ala Val Ser Lys Phe Glu Lys Asp Ser
89          100                105                110
90      Ser Ser Ser Ser Ser Ser Asp Ser Ser Thr Lys Pro Glu Ala Ser Asp
91          115                120                125
92      Thr Ala Lys Pro Asn Lys Pro Thr Glu Pro Gly Glu Lys Val Ala Glu
93          130                135                140
94      Ala Lys Lys Lys Val Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys

```

PAGE: 3

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/056,019A

DATE: 07/07/1999
TIME: 14:28:17

Input Set: I056019A.RAW

95	145	150	155	160
96	Glu Glu Asp Arg Arg Asn Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Glu			
97		165	170	175
98	Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala Glu Leu Glu			
99		180	185	190
100	Leu Val Lys Val Lys Ala Asn Glu Pro Arg Asp Glu Gln Lys Ile Lys			
101		195	200	205
102	Gln Ala Glu Ala Glu Val Glu Ser Lys Gln Ala Glu Ala Thr Arg Leu			
103		210	215	220
104	Lys Lys Ile Lys Thr Asp Arg Glu Glu Ala Glu Glu Glu Ala Lys Arg			
105	225		230	235
106	Arg Ala Asp Ala Lys Glu Gln Gly Lys Pro Lys Gly Arg Ala Lys Arg			
107		245	250	255
108	Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala			
109		260	265	270
110	Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Pro Ser Pro Ser			
111		275	280	285
112	Leu Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu			
113		290	295	300
114	Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr			
115	305		310	315
116	Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp			
117		325	330	335
118	Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys			
119		340	345	350
120	Glu Pro Arg Asn Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu			
121		355	360	365
122	Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg			
123		370	375	380
124	Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Ala Glu Glu Asp Lys			
125	385		390	395
126	Val Lys Glu Lys Pro Ala Glu Gln Pro Gln Pro Ala Pro Ala Pro Lys			
127		405	410	415
128	Ala Glu Lys Pro Ala Pro Ala Pro Lys Pro Glu Asn Pro Ala Glu Gln			
129		420	425	430
130	Pro Lys Ala Glu Lys Pro Ala Asp Gln Gln Ala Glu Glu Asp Tyr Ala			
131		435	440	445
132	Arg Arg Ser Glu Glu Glu Tyr Asn Arg Leu Thr Gln Gln Gln Pro Pro			
133		450	455	460
134	Lys Thr Glu Lys Pro Ala Gln Pro Ser Thr Pro Lys Thr Gly Trp Lys			
135	465		470	475
136	Gln Glu Asn Gly Met Trp Tyr Phe Tyr Asn Thr Asp Gly Ser Met Ala			
137		485	490	495
138	Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr Tyr Leu Asn Ser Asn			
139		500	505	510
140	Gly Ala Met Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr Tyr			
141		515	520	525
142	Leu Asn Ala Asn Gly Ser Met Ala Thr Gly Trp Leu Gln Asn Asn Gly			
143		530	535	540
144	Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ser Met Ala Thr Gly Trp Leu			

PAGE: 4

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/056,019A

DATE: 07/07/1999
TIME: 14:28:17

Input Set: I056019A.RAW

```

145      545      550      555      560
146      Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ser Met Ala
147      565      570      575
148      Thr Gly Trp Leu Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn
149      580      585      590
150      Gly Asp Met Ala Thr Gly Trp Val Lys Asp Gly Asp Thr Trp Tyr Tyr
151      595      600      605
152      Leu Glu Ala Ser Gly Ala Met Lys Ala Ser Gln Trp Phe Lys Val Ser
153      610      615      620
154      Asp Lys Trp Tyr Tyr Val Asn Gly Ser Gly Ala Leu Ala Val Asn Thr
155      625      630      635      640
156      Thr Val Asp Gly Tyr Gly Val Asn Ala Asn Gly Glu Trp Val Asn
157      645      650      655
158 <210> SEQ ID NO 3
159 <211> LENGTH: 284
160 <212> TYPE: PRT
161 <213> ORGANISM: Streptococcus pneumoniae
162 <400> SEQUENCE: 3
163      Glu Asn Glu Gly Ala Thr Gln Val Pro Thr Ser Ser Asn Arg Ala Asn
164      1      5      10      15
165      Glu Ser Gln Ala Glu Gln Gly Glu Gln Pro Lys Lys Leu Asp Ser Glu
166      20      25      30
167      Arg Asp Lys Ala Arg Lys Glu Val Glu Glu Tyr Val Lys Lys Ile Val
168      35      40      45
169      Gly Glu Ser Tyr Ala Lys Ser Thr Lys Lys Arg His Thr Ile Thr Val
170      50      55      60
171      Ala Leu Val Asn Glu Leu Asn Asn Ile Lys Asn Glu Tyr Leu Asn Lys
172      65      70      75      80
173      Ile Val Glu Ser Thr Ser Glu Ser Gln Leu Gln Ile Leu Met Met Glu
174      85      90      95
175      Ser Arg Ser Lys Val Asp Glu Ala Val Ser Lys Phe Glu Lys Asp Ser
176      100      105      110
177      Ser Ser Ser Ser Ser Ser Asp Ser Ser Thr Lys Pro Glu Ala Ser Asp
178      115      120      125
179      Thr Ala Lys Pro Asn Lys Pro Thr Glu Pro Gly Glu Lys Val Ala Glu
180      130      135      140
181      Ala Lys Lys Lys Val Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys
182      145      150      155      160
183      Glu Glu Asp Arg Arg Asn Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Glu
184      165      170      175
185      Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala Glu Leu Glu
186      180      185      190
187      Leu Val Lys Val Lys Ala Asn Glu Pro Arg Asp Glu Gln Lys Ile Lys
188      195      200      205
189      Gln Ala Glu Ala Glu Val Glu Ser Lys Gln Ala Glu Ala Thr Arg Leu
190      210      215      220
191      Lys Lys Ile Lys Thr Asp Arg Glu Glu Ala Glu Glu Glu Ala Lys Arg
192      225      230      235      240
193      Arg Ala Asp Ala Lys Glu Gln Gly Lys Pro Lys Gly Arg Ala Lys Arg
194      245      250      255

```

PAGE: 5

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/056,019A

DATE: 07/07/1999
TIME: 14:28:17

Input Set: I056019A.RAW

```

195      Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala
196                      260                      265                      270
197      Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu
198                      275                      280
199 <210> SEQ ID NO 4
200 <211> LENGTH: 106
201 <212> TYPE: PRT
202 <213> ORGANISM: Streptococcus pneumoniae
203 <400> SEQUENCE: 4
204      Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala
205      1                      5                      10                      15
206      Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro
207                      20                      25                      30
208      Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp Val
209                      35                      40                      45
210      Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu
211      50                      55                      60
212      Pro Arg Asn Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu Ser
213      65                      70                      75                      80
214      Lys Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg Lys
215                      85                      90                      95
216      Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala
217                      100                      105
218 <210> SEQ ID NO 5
219 <211> LENGTH: 109
220 <212> TYPE: PRT
221 <213> ORGANISM: Streptococcus pneumoniae
222 <400> SEQUENCE: 5
223      Thr Glu Pro Gly Glu Lys Val Ala Glu Ala Lys Lys Lys Val Glu Glu
224      1                      5                      10                      15
225      Ala Glu Lys Lys Ala Lys Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr
226                      20                      25                      30
227      Pro Thr Ile Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp
228                      35                      40                      45
229      Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Val Lys Ala Asn
230      50                      55                      60
231      Glu Pro Arg Asp Glu Gln Lys Ile Lys Gln Ala Glu Ala Glu Val Glu
232      65                      70                      75                      80
233      Ser Lys Gln Ala Glu Ala Thr Arg Leu Lys Lys Ile Lys Thr Asp Arg
234                      85                      90                      95
235      Glu Glu Ala Glu Glu Glu Ala Lys Arg Arg Ala Asp Ala
236                      100                      105
237 <210> SEQ ID NO 6
238 <211> LENGTH: 4
239 <212> TYPE: PRT
240 <213> ORGANISM: Streptococcus pneumoniae
241 <220> FEATURE:
242 <221> NAME/KEY: NON_CONS
243 <222> LOCATION: (2)..(3)
244 <223> OTHER INFORMATION: They could be any amino acid at these two

```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY
PATENT APPLICATION US/09/056,019ADATE: 07/07/1999
TIME: 14:28:17

Input Set: I056019A.RAW

Line	? Error/Warning	Original Text
247	W "N" or "Xaa" used: Feature required	Lys Xaa Xaa Glu
797	W "N" or "Xaa" used: Feature required	Xaa Glu Asn Glu Gly
838	W "N" or "Xaa" used: Feature required	Glu Val Xaa Gly Glu Leu Ala Thr Pro Asp L